

An appeal for declaring research goals

Brandon Rohrer

*Sandia National Laboratories
Albuquerque, NM 87185, USA*

BROHRER@GMAIL.COM

Editor: Wlodzislaw Duch

In their review, the authors present a challenge task, exploration of the natural universe, and assert that performing it would fulfill the main goals of AGI. From this challenge, they distill four sub-goals that they consider necessary to achieve it—realtime, learning, resource management, and meta-learning—and use these as criteria for comparing several cognitive architectures.¹ The resulting autonomy-centered comparison provides valuable insights into the relative capabilities of different types of architectures. However, what they do not mention is that autonomy is not the only motivator for creators of cognitive architectures. The explicit goal of a number of architectures is to model the operation of the human brain.

In an informal show-of-hands survey at the 2009 Biologically Inspired Cognitive Architectures Symposium, the audience (consisting primarily of cognitive architecture creators) was asked to choose one of two hypothetical outcomes for their research: 1) Create a system that equaled human performance in all areas, but provided no insights into the operation of the brain, and 2) Create a system that demonstrably modeled every aspect of the brain, but could do nothing. Approximately half the respondents chose each option. Informal communications with a number of researchers since then have supported these results. Within the AGI community, there are strong emphases on both modeling and autonomy.

While this was a patently non-rigorous survey, not to mention a false dichotomy, it does illustrate the range of motivators in creating cognitive architectures. In fact, detailed conversations with practitioners support the notion that there are as many motivators for creating AGIs (and thus as many definitions of “success”) as there are researchers. Comparing architectures on any one measure is instructive, but partial. This situation can be frustrating. With such orthogonal goals, modelers and autonomists often have difficulty communicating. And without explicit goals, presentations of model-focused architectures resemble beauty contests, and demonstrations of autonomy-focused architectures play like a diverting series of pet tricks.

One way to address this is to be painfully explicit about our own goals. If we can definitively answer the questions “**What are you trying to do?**” and “**How do you know if you’re getting better at it?**” we give clarity to our work. The answers are still useful, even if they change over time and from researcher to researcher. They provide a measure of progress by allowing us to compare our work to each other and to our past selves.

1. Disclaimer: I am also the author of a cognitive architecture, BECCA. Like the authors, my own goals are strongly rooted in autonomy, focused on natural world interaction, (Rohrer, 2010) that is, creating a system that can do everything I can do. The intermediate tasks that I’m working toward are detailed in (Rohrer, 2010), the benchmark I am currently testing BECCA against is described in (Rohrer, 2012), and its Python code is downloadable from (Chapman, 2012).

The more specific we can be in our goals, the more they will benefit us. For modeling-focused architectures, What is the scope of your modeling? Which phenomena are you trying to model? Which data sets will you compare your performance to? For autonomy-focused architectures, What task space are you planning to operate in? Which tasks are you trying to perform? How will you measure your performance on them? If we can represent our architectures' fitness with a number, we can easily demonstrate their improvement over time.

It is true that clarity exposes us to some risk. After all, what if we fail to make progress? Or what if someone outperforms us on our own tasks? But the benefits of clarity are too great to ignore. Clear goals keep our research focused, help us to allocate scarce resources to maximum effect, and simplify our communication. Being able to directly compare cognitive architectures provides a coherence to the field that increases its perceived rigor and legitimacy. And perhaps most importantly, specific goals help us to concisely describe our vision, helping us to capture the imagination of a new generation of researchers.

Acknowledgments

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

- Chapman, M. 2012. BECCA home page. <http://www.openbecca.org>.
- Rohrer, B. 2010. Accelerating progress in Artificial General Intelligence: Choosing a benchmark for natural world interaction. *Journal of Artificial General Intelligence* 2:1–28.
- Rohrer, B. 2012. BECCA 0.4.0 User's Guide.